

INVARIANT GALTON–WATSON TREES: METRIC PROPERTIES AND ATTRACTION WITH RESPECT TO GENERALIZED DYNAMICAL PRUNING

YEVGENIY KOVCHEGOV * ** AND
GUOCHEN XU,* *** Oregon State University
ILYA ZALIAPIN,**** University of Nevada Reno

Abstract

The *invariant Galton–Watson (IGW)* tree measures are a one-parameter family of critical Galton–Watson measures invariant with respect to a large class of tree reduction operations. Such operations include the generalized dynamical pruning (also known as hereditary reduction in a real tree setting) that eliminates descendant subtrees according to the value of an arbitrary subtree function that is monotone nondecreasing with respect to an isometry-induced partial tree order. We show that, under a mild regularity condition, the IGW measures are attractors of arbitrary critical Galton–Watson measures with respect to the generalized dynamical pruning. We also derive the distributions of height, length, and size of the IGW trees.

Keywords: Critical branching process; random tree; attractor

2020 Mathematics Subject Classification: Primary 60J80
Secondary 60G18; 35B41

1. Introduction

In this paper we continue the study of the critical branching processes with the progeny generating function $Q(z) = z + q(1 - z)^{1/q}$ for a given parameter $q \in [1/2, 1)$. The importance of these processes was previously observed in [1, 7, 9, 20, 21, 23, 25, 32]. The random tree measures induced by these branching processes are called here the *invariant Galton–Watson (IGW)* measures. This paper has two goals. First, we establish the main metric properties of the IGW trees: the distributions of their height, length, and size (the number of edges). These distributions are well-studied for a special case of the IGW process with $q = 1/2$ that coincides with the critical binary Galton–Watson process [20]. Here we establish the results for the general case of $q \in [1/2, 1)$. Second, we extend the results of [21], where the IGW trees were shown to be the attractors of the pushforward measures under the iterative application of Horton pruning (eliminating tree leaves followed by a series reduction). Here, we obtain analogous results under a much broader *generalized dynamical pruning* introduced in [19].

Received 25 January 2022; revision received 28 June 2022.

* Postal address: Department of Mathematics, Oregon State University, Corvallis, OR 97331-4605, USA.

** Email address: yevgeniy.kovchegov@oregonstate.edu

*** Email address: xugu@oregonstate.edu

**** Postal address: Department of Mathematics and Statistics, University of Nevada Reno, Reno, NV 89557-0172, USA. Email address: zal@unr.edu

© The Author(s), 2023. Published by Cambridge University Press on behalf of Applied Probability Trust.

Since the generalized dynamical pruning can be expressed via the *hereditary reduction* of [9], the attractor property of IGW tree measures holds for the hereditary reduction as well. Also, the IGW trees turn out to be the attractors under the *Bernoulli leaf coloring*, a tree reduction studied in [8].

The paper is organized as follows. Section 2 contains all the necessary background. The results are stated in Section 3 and proved in Section 4. The paper concludes with a discussion in Section 5. Appendix A contains the statement of the Lagrange inversion theorem used in this paper. In Appendix B, Karamata's theorem and its converse are stated. Appendix C contains a proof of Lemma 2.5.

2. Background

2.1. Spaces of trees

A tree is called *rooted* if one of its vertices, denoted by ρ , is designated as the tree root. The existence of root imposes a parent–offspring relation between each pair of adjacent vertices: the one closer to the root is called the *parent*, and the other the *offspring*. A tree is called *reduced* if it has no vertices of degree 2, with the root as the only possible exception. Let \mathcal{T} denote the space of finite unlabeled rooted reduced trees with no planar embedding. The absence of planar embedding is the absence of order among the offspring of the same parent. The space \mathcal{T} includes the *empty tree* ϕ comprising a root vertex ρ and no edges.

Let \mathcal{L} denote the space of trees from \mathcal{T} equipped with edge lengths. Thus, a tree in \mathcal{L} is itself a metric space. A metric tree $T \in \mathcal{L}$ can be considered as a metric space with distance $d(\cdot, \cdot)$ induced by the Lebesgue measure along the tree edges [20]. Hence, a metric tree $T \in \mathcal{L}$ can be represented as a pair $T = (S, d)$, where S represents the space and d is the metric defined on the space S . The operator $\text{SHAPE}(T) : \mathcal{L} \rightarrow \mathcal{T}$ projects a tree $T \in \mathcal{L}$ with edge lengths on a tree in \mathcal{T} that retains the combinatorial structure of T (and drops the edge length assignment).

A nonempty rooted tree is called *planted* if its root ρ has degree one. In this case the only edge connected to the root is called the *stem*. If the root ρ is of degree ≥ 2 then the tree is called *stemless*. We denote by \mathcal{T}^\dagger and \mathcal{T}^\vee the subspaces of \mathcal{T} consisting of planted and stemless trees, respectively. Similarly, \mathcal{L}^\dagger and \mathcal{L}^\vee are the subspaces of \mathcal{L} consisting of planted and stemless trees. Additionally, we include the empty tree $\phi = \{\rho\}$ as an element in each of these subspaces, \mathcal{T}^\dagger , \mathcal{T}^\vee , \mathcal{L}^\dagger , and \mathcal{L}^\vee , defined above.

2.2. Galton–Watson tree measures

Consider a Galton–Watson branching process with a given progeny distribution (probability mass function) $\{q_k\}$, $k = 0, 1, 2, \dots$. More specifically, we consider a discrete-time Markov process that begins with a single progenitor, which produces a single offspring (hence the trees we examine are planted). At each later step, each existing population member produces $k \geq 0$ offspring with probability q_k , independently from the prior history of the process; see [3, 13].

A *Galton–Watson tree* is formed by the trajectory of the Galton–Watson branching process, with the progenitor corresponding to the tree root ρ . The single offspring of the progenitor is represented in the tree by the vertex connected to the tree root by the stem [20]. We denote by $\mathcal{GW}(\{q_k\})$ the probability measure on \mathcal{T}^\dagger induced by the Galton–Watson process with progeny distribution $\{q_k\}$. Assuming $q_1 < 1$, the resulting tree is finite with probability one if and only if $\sum_{k=0}^{\infty} kq_k \leq 1$, i.e., the Galton–Watson process is either subcritical or critical. In this paper we let $q_1 = 0$ so that the Galton–Watson process generates a *reduced* tree.

For a given probability mass function $\{q_k\}$ with $q_1 = 0$ and a positive real λ , consider a random tree T in \mathcal{L}^1 satisfying $\text{SHAPE}(T) \stackrel{d}{\sim} \mathcal{GW}(\{q_k\})$, and such that, conditioned on $\text{SHAPE}(T)$, the edge lengths are distributed as independent and identically distributed (i.i.d.) exponential random variables with parameter λ . Let $\mathcal{GW}(\{q_k\}, \lambda)$ denote the distribution of the random tree T thus defined. Measures $\mathcal{GW}(\{q_k\})$ and $\mathcal{GW}(\{q_k\}, \lambda)$ induced by critical (or subcritical) branching processes will be called *critical (or subcritical) Galton–Watson measures*.

2.3. Invariant Galton–Watson measures

The invariant Galton–Watson measures form a single-parameter family of critical Galton–Watson measures $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$ on \mathcal{T}^1 , which we define as follows.

Definition 2.1. (*Invariant Galton–Watson measures in \mathcal{T}^1*) For a given $q \in [1/2, 1)$, a critical Galton–Watson measure $\mathcal{GW}(\{q_k\})$ on \mathcal{T}^1 is said to be the *invariant Galton–Watson (IGW) measure* with parameter q , and denoted by $\mathcal{IGW}(q)$, if its generating function is given by

$$Q(z) = z + q(1 - z)^{1/q}. \tag{1}$$

The respective branching probabilities are $q_0 = q, q_1 = 0, q_2 = (1 - q)/2q$, and

$$q_k = \frac{1 - q}{k! q} \prod_{i=2}^{k-1} (i - 1/q) \quad \text{for } k \geq 3. \tag{2}$$

Here, if $q = 1/2$, then the distribution is critical binary, i.e., $\mathcal{GW}(q_0=q_2=1/2)$. If $q \in (1/2, 1)$, the distribution is of Zipf type with

$$q_k = \frac{(1 - q)\Gamma(k - 1/q)}{q\Gamma(2 - 1/q)k!} \sim Ck^{-(1+q)/q}, \quad \text{where } C = \frac{1 - q}{q\Gamma(2 - 1/q)}. \tag{3}$$

We notice that

$$Q'(z) = 1 - (1 - z)^{1/q-1}, \quad Q''(z) = (1 - z)^{1/q-2},$$

which implies that $Q'(1) = 1$, that is, the IGW processes are critical, and $Q''(1) < \infty$, that is, the offspring distribution’s second moment is finite, if and only if $q = 1/2$.

These tree measures (Definition 2.1) are also known as *stable Galton–Watson trees* or Galton–Watson trees with *stable offspring distribution* [9]. They have previously been considered in the work of Zolotarev [32], Neveu [25], Le Jan [23], Duquesne and Le Gall [7], Abraham and Delmas [1], and Duquesne and Winkel [9]. Moreover, in [25], Neveu regards the generating functions (1) to be the most important in the critical case.

The definition of the IGW measure can be extended to \mathcal{L}^1 by assigning i.i.d. exponentially distributed edge lengths.

Definition 2.2. (*Invariant Galton–Watson measures in \mathcal{L}^1*) For a given $q \in [1/2, 1)$ and $\lambda > 0$, a random tree T in \mathcal{L}^1 is said to be an *exponential invariant Galton–Watson tree* if it satisfies the following properties:

- (i) $\text{SHAPE}(T) \stackrel{d}{\sim} \mathcal{IGW}(q)$;
- (ii) conditioned on $\text{SHAPE}(T)$, the edge lengths are distributed as i.i.d. exponential random variables with parameter $\lambda > 0$.

Such a tree is denoted by $\mathcal{IGW}(q, \lambda)$. In other words, $T \stackrel{d}{\sim} \mathcal{GW}(\{q_k\}, \lambda)$ with q_k as in (2).

2.4. Invariance and attractor properties of IGW family under Horton pruning

Horton pruning of a tree T (in \mathcal{T} or \mathcal{L}) is done by removing all the leaves of T (leaf vertices together with the corresponding adjacent edges) followed by consecutive series reduction (removing degree-two vertices by merging adjacent edges into one and adding up their lengths for a tree in \mathcal{L}). The resulting reduced tree is denoted by $\mathcal{R}(T)$. We refer to [20] for a detailed treatment of Horton pruning. The Horton pruning operator \mathcal{R} induces a map on \mathcal{T} (or \mathcal{L}). The trajectory of each tree T under iterative application of \mathcal{R} , i.e.,

$$T \equiv \mathcal{R}^0(T) \rightarrow \mathcal{R}^1(T) \rightarrow \dots \rightarrow \mathcal{R}^k(T) = \phi, \tag{4}$$

is uniquely defined and finite with the empty tree ϕ as the (only) fixed point. In [19], it was established that, under iterative Horton pruning, the $\mathcal{IGW}(q)$ measures are the attractors of all critical Galton–Watson trees that satisfy the following regularity assumption.

Assumption 2.1. *Consider a critical Galton–Watson measure $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$ and the respective progeny generating function $Q(z)$. We assume that the following limit exists:*

$$\lim_{x \rightarrow 1^-} \frac{Q(x) - x}{(1-x)(1-Q'(x))}. \tag{5}$$

We will use the function $g(x)$ defined in the following proposition from [21].

Proposition 2.1. ([21].) *Consider a critical Galton–Watson measure $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$ and the respective progeny generating function $Q(z)$. Then*

$$Q(x) - x = (1-x)^2 g(x),$$

where $g(x)$ is defined as follows. Let $X \stackrel{d}{\sim} \{q_k\}$ be a progeny random variable; then

$$g(x) = \sum_{m=0}^{\infty} \mathbb{E}[(X - m - 1)_+] x^m = \sum_{m=0}^{\infty} \sum_{k=m+1}^{\infty} (k - m - 1) q_k x^m, \tag{6}$$

where $x_+ = \max\{x, 0\}$.

An important limit is defined in the following lemma.

Lemma 2.1. ([21].) *Consider a critical Galton–Watson measure $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$. If Assumption 2.1 is satisfied, then for $g(x)$ defined in (6) the limit*

$$\lim_{x \rightarrow 1^-} \left(\frac{\ln g(x)}{-\ln(1-x)} \right) = L \tag{7}$$

exists, and

$$\lim_{x \rightarrow 1^-} \frac{Q(x) - x}{(1-x)(1-Q'(x))} = \frac{1}{2-L}.$$

The following three results (Lemmas 2.2, 2.3, and 2.4) concerning the applicability of Assumption 2.1 and the limit L in (7) were established in [21].

Lemma 2.2. ([21].) *Consider a critical Galton–Watson measure $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$. For a progeny variable $X \stackrel{d}{\sim} \{q_k\}$ and $g(x)$ in (6), if*

$$\mathbb{E}[X^{2-\epsilon}] = \sum_{k=0}^{\infty} k^{2-\epsilon} q_k < \infty \quad \forall \epsilon > 0, \tag{8}$$

then $L = \lim_{x \rightarrow 1^-} \left(\frac{\ln g(x)}{-\ln(1-x)} \right) = 0$. If moreover the second moment is finite, i.e.,

$$\mathbb{E}[X^2] = \sum_{k=0}^{\infty} k^2 q_k < \infty,$$

then Assumption 2.1 is satisfied with $\lim_{x \rightarrow 1^-} \frac{Q(x)-x}{(1-x)(1-Q'(x))} = \frac{1}{2}$.

The next lemma provides a basic regularity condition for Assumption 2.1 to hold.

Lemma 2.3. (Regularity condition [21].) Consider a critical Galton–Watson measure $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$ and infinite second moment, i.e., $\sum_{k=0}^{\infty} k^2 q_k = \infty$. Suppose that for the progeny variable $X \stackrel{d}{\sim} \{q_k\}$ the following limit exists:

$$\Lambda = \lim_{k \rightarrow \infty} \frac{k}{\mathbb{E}[X | X \geq k]} = \lim_{k \rightarrow \infty} \frac{k \sum_{m=k}^{\infty} q_m}{\sum_{m=k}^{\infty} m q_m}. \tag{9}$$

Then Assumption 2.1 is satisfied with $\lim_{x \rightarrow 1^-} \frac{Q(x)-x}{(1-x)(1-Q'(x))} = 1 - \Lambda$ and $L = 2 + \frac{1}{1-\Lambda}$.

The next lemma follows immediately from Lemma 2.3.

Lemma 2.4. (Zipf distribution [21].) Consider a critical Galton–Watson process $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$ and offspring distribution $\{q_k\}$ of Zipf type:

$$q_k \sim Ck^{-(\alpha+1)} \quad \text{with } \alpha \in (1, 2] \text{ and } C > 0. \tag{10}$$

Then Assumption 2.1 is satisfied, and

$$L = \lim_{x \rightarrow 1^-} \left(\frac{\ln g(x)}{-\ln(1-x)} \right) = 2 - \alpha. \tag{11}$$

In Sections 3.2 and 3.3 of this paper we consider generalizations of the following two theorems, which were proved in [21].

Theorem 2.1. (Self-similarity under Horton pruning [21].) Consider a critical or subcritical Galton–Watson measure $\mu \equiv \mathcal{GW}(\{q_k\})$ with $q_1 = 0$ that satisfies Assumption 2.1. Then a Galton–Watson measure μ is Horton-prune-invariant (self-similar), i.e., the pushforward measure $\nu(T) = \mu \circ \mathcal{R}^{-1}(T) = \mu(\mathcal{R}^{-1}(T))$ satisfies $\nu(T | T \neq \emptyset) = \mu(T)$, if and only if μ is the IGW measure $\mathcal{IGW}(q)$ with $q \in [1/2, 1)$.

Theorem 2.2. (IGW attractors under iterative Horton pruning [21].) Consider a critical Galton–Watson measure $\rho_0 \equiv \mathcal{GW}(\{q_k\})$ with $q_1 = 0$ on \mathcal{T}^1 . Starting with $k = 0$, and for each consecutive integer, let $\nu_k = \mathcal{R}_*(\rho_k)$ denote the pushforward probability measure induced by the pruning operator, i.e., $\nu_k(T) = \rho_k \circ \mathcal{R}^{-1}(T) = \rho_k(\mathcal{R}^{-1}(T))$, and set $\rho_{k+1}(T) = \nu_k(T | T \neq \emptyset)$. Suppose Assumption 2.1 is satisfied. Then, for any $T \in \mathcal{T}^1$,

$$\lim_{k \rightarrow \infty} \rho_k(T) = \rho^*(T),$$

where ρ^* denotes the IGW measure $\mathcal{IGW}(q)$ with $q = \frac{1}{2-L}$ and L as defined in (7).

Finally, if the Galton–Watson measure $\rho_0 \equiv \mathcal{GW}(\{q_k\})$ is subcritical, then $\rho_k(T)$ converges to a point mass measure, $\mathcal{GW}(q_0=1)$.

2.5. Generalized dynamical pruning

Given a metric tree $T = (S, d) \in \mathcal{L}$ and a point $x \in S$, let $\Delta_{x,T}$ be the *descendant tree* of x : the tree comprising all points of T descendant to x , including x . Then $\Delta_{x,T}$ is itself a tree in \mathcal{L} with the root at x .

Let $T_1 = (S_1, d_1)$ and $T_2 = (S_2, d_2)$ be two metric rooted trees, and let ρ_1 denote the root of T_1 . A function $f: T_1 \rightarrow T_2$ is said to be an *isometry* if $\text{Image}[f] \subseteq \Delta_{f(\rho_1), T_2}$ and for all pairs $x, y \in T_1$,

$$d_2(f(x), f(y)) = d_1(x, y).$$

We use the above-defined notion of isometry to define a *partial order* in the space \mathcal{L} as follows. We say that T_1 is *less than or equal to* T_2 , and write $T_1 \preceq T_2$, if there is an isometry $f: T_1 \rightarrow T_2$. The relation \preceq is a partial order as it satisfies the reflexivity, antisymmetry, and transitivity conditions. We say that a function $\varphi: \mathcal{L} \rightarrow \mathbb{R}$ is *monotone nondecreasing* with respect to the partial order \preceq if $\varphi(T_1) \leq \varphi(T_2)$ whenever $T_1 \preceq T_2$.

Next, we recall the definition of the *generalized dynamical pruning* as stated in [19, 20]. Consider a *monotone nondecreasing* function $\varphi: \mathcal{L} \rightarrow \mathbb{R}_+$ with respect to the above-defined partial order \preceq . We define the generalized dynamical pruning operator $\mathcal{S}_t(\varphi, T): \mathcal{L} \rightarrow \mathcal{L}$ induced by φ for any given time parameter $t \geq 0$ as

$$\mathcal{S}_t(\varphi, T) := \{\rho\} \cup \left\{ x \in T \setminus \{\rho\} : \varphi(\Delta_{x,T}) \geq t \right\}, \quad (12)$$

where ρ denotes the root of the tree T . Informally, the operator \mathcal{S}_t cuts all subtrees $\Delta_{x,T}$ for which the value of φ is below the threshold t , and always keeps the tree root.

Below we discuss some well-studied examples of generalized dynamical pruning.

Example 2.1. (*Pruning via the Horton–Strahler order*) The Horton–Strahler order [5, 17, 20, 29] was initially introduced in the context of geomorphology. It can be defined via the operation of Horton pruning \mathcal{R} . The Horton–Strahler order $\text{ord}(T)$ of a planted tree from \mathcal{L}^1 (or \mathcal{T}^1) is the minimal number of prunings necessary to eliminate a tree T . The Horton–Strahler order $\text{ord}(T)$ of a stemless tree from \mathcal{L}^\vee (or \mathcal{T}^\vee) equals one plus the minimal number of prunings necessary to eliminate a tree T . For a tree T in either \mathcal{T} or \mathcal{L} , consider

$$\varphi(T) = \text{ord}(T) - 1. \quad (13)$$

For $k \in \mathbb{N}$, let \mathcal{R}^k denote the k th iteration of Horton pruning \mathcal{R} , i.e., $\mathcal{R}^0(T) = T$ and $\mathcal{R}^k = \underbrace{\mathcal{R} \circ \dots \circ \mathcal{R}}_{k \text{ times}}$. With the function φ as in (13), the generalized dynamical pruning operator $\mathcal{S}_t = \mathcal{R}^{\lfloor t \rfloor}$ satisfies the *discrete semigroup property* [19, 20]:

$$\mathcal{S}_t \circ \mathcal{S}_s = \mathcal{S}_{t+s} \text{ for any } t, s \in \mathbb{N}_0,$$

as $\mathcal{R}^t \circ \mathcal{R}^s = \mathcal{R}^{t+s}$. A recent survey of results related to invariance of a tree distribution with respect to Horton pruning is given in [20].

Example 2.2. (*Pruning via the tree height*) If we let the $\varphi(T)$ be the height function, i.e., for a tree $T \in \mathcal{L}$, let

$$\varphi(T) = \text{HEIGHT}(T), \quad (14)$$

then the generalized dynamical pruning $\mathcal{S}_t(\cdot) = \mathcal{S}_t(\varphi, \cdot)$ will coincide with the continuous pruning (leaf-length erasure) studied in Neveu [25], which established the invariance of critical binary Galton–Watson measures with i.i.d. exponential edge lengths with respect to this

operation. In this case the operator \mathcal{S}_t is known to satisfy the continuous semigroup property [9, 19, 25]:

$$\mathcal{S}_t \circ \mathcal{S}_s = \mathcal{S}_{t+s} \text{ for any } t, s \geq 0.$$

Example 2.3. (*Pruning via the tree length*) Let the function $\varphi(T)$ equal the total length of $T \in \mathcal{L}$:

$$\varphi(T) = \text{LENGTH}(T). \tag{15}$$

The pruning operator $\mathcal{S}_t(\cdot) = \mathcal{S}_t(\varphi, \cdot)$ with the pruning function φ as in (15) coincides with the potential dynamics of the continuum mechanics formulation of the one-dimensional ballistic annihilation model $A + A \rightarrow \emptyset$ [19]. Importantly, the operator \mathcal{S}_t induced by the length function φ as in (15) does not satisfy the semigroup property (discrete or continuous); i.e., $\mathcal{S}_t \circ \mathcal{S}_s \neq \mathcal{S}_{t+s}$ [19].

Example 2.4. (*Pruning via the number of leaves*) Let $\text{LEAVES}(T)$ denote the number of leaves in a tree T . Then

$$\varphi(T) = \text{LEAVES}(T) \tag{16}$$

is another monotone nondecreasing function. The generalized dynamical pruning operator $\mathcal{S}_t(\cdot) = \mathcal{S}_t(\varphi, \cdot)$ induced by φ as in (16) does not satisfy the semigroup property, whether discrete or continuous. This type of pruning naturally arises in the context of Shreve stream ordering in hydrodynamics.

2.6. Generalized dynamical pruning as a hereditary reduction

Duquesne and Winkel [9] introduced a very general kind of tree reduction in the context of *complete locally compact rooted (CLCR) real trees*, which include all the trees in \mathcal{L} . In [9], a *hereditary property* A is defined as a Borel subset in the space \mathbb{T} of CLCR real trees (more precisely, their equivalence classes under isometry) equipped with the pointed Gromov–Hausdorff metric such that for a CLCR real tree $T \in \mathbb{T}$ and any $x \in T$,

$$\Delta_{x,T} \in A \implies T = \Delta_{\rho,T} \in A.$$

As an example of a hereditary property, one may consider $A = \{T \in \mathbb{T} : \text{HEIGHT}(T) \geq t\}$.

A hereditary property $A \subset \mathbb{T}$ induces a *hereditary reduction* operator $R_A : \mathbb{T} \rightarrow \mathbb{T}$ defined as

$$R_A(T) := \{\rho\} \cup \{x \in T \setminus \rho : \Delta_{x,T} \in A\}. \tag{17}$$

The following result was proved in [9, Theorem 2.18].

Theorem 2.3. (Evolution of Galton–Watson trees under hereditary reduction [9].) *Consider a critical or subcritical Galton–Watson measure $\mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$ ($q_1 = 0$) on \mathcal{L}^1 with generating function $Q(z)$. For a hereditary property $A \subset \mathbb{T}$, let ν denote the corresponding pushforward probability measure induced by the hereditary reduction R_A :*

$$\nu(T) = \mu \circ R_A^{-1}(T) = \mu(R_A^{-1}(T)).$$

Then $\nu(T \in \cdot \mid R_A(T) \neq \phi) \stackrel{d}{=} \mathcal{GW}(\{g_k\}, \lambda(1 - Q'(1 - p)))$ is a Galton–Watson tree measure over \mathcal{L}^1 with independent exponential edge lengths with parameter $\lambda(1 - Q'(1 - p))$, and generating function

$$G(z) = z + \frac{Q((1 - p) + pz) - (1 - p) - pz}{p(1 - Q'(1 - p))}, \tag{18}$$

where $p = \mathbb{P}(R_A(T) \neq \phi)$.

Observe the following direct link between the operations of generalized dynamical pruning and hereditary reduction. Consider a Borel measurable monotone nondecreasing function $\varphi : \mathcal{L} \rightarrow \mathbb{R}_+$. Then, for a fixed $t \geq 0$, the Borel set

$$A = \{T \in \mathbb{T} : \varphi(T) \geq t\} \tag{19}$$

is a hereditary property, and therefore $\mathcal{S}_t(\varphi, T) = R_A(T)$ is a hereditary reduction.

The composition of two hereditary properties A and A' was defined in [9, Definition 2.12] as the set

$$A' \circ A = \{T \in \mathbb{T} : R_A(T) \in A'\}.$$

Consequently, in Lemma 2.13 of [9], the hereditary reductions were shown to satisfy the composition property, $R_{A' \circ A} = R_{A'} \circ R_A$. Importantly, if we let A_t denote the hereditary property in (19), then

$$A_t \circ A_s \neq A_{s+t}$$

for many (or rather, all but a few) functions φ , e.g. for $\varphi(T) = \text{LENGTH}(T)$. Speaking of the exceptions, the equation $A_t \circ A_s = A_{s+t}$ is known to hold for $\varphi(T) = \text{HEIGHT}(T)$ with real $s, t \in [0, \infty)$, corresponding to Neveu (leaf-length) erasure as in Example 2.2, and for $\varphi(T) = \text{ord}(T) - 1$ with integer $s, t \in \mathbb{Z}_+$, corresponding to Horton pruning as in Example 2.1.

We will need the following adaptation of Theorem 2.3 for generalized dynamical pruning.

Lemma 2.5. (Pruning Galton–Watson trees) *Consider a critical or subcritical Galton–Watson measure $\mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$ ($q_1 = 0$) on \mathcal{L}^1 with generating function $Q(z)$. For a monotone nondecreasing function $\varphi : \mathcal{L} \rightarrow \mathbb{R}_+$, let ν denote the corresponding pushforward probability measure induced by the pruning operator $\mathcal{S}_t(\varphi, T) = \mathcal{S}_t(\varphi, T)$:*

$$\nu(T) = \mu \circ \mathcal{S}_t^{-1}(T) = \mu(\mathcal{S}_t^{-1}(T)).$$

Then $\nu(T \in \cdot \mid T \neq \phi) \stackrel{d}{=} \mathcal{GW}(\{g_k\}, \lambda(1 - Q'(1 - p_t)))$ is a Galton–Watson tree measure over \mathcal{L}^1 with independent exponential edge lengths with parameter $\lambda(1 - Q'(1 - p_t))$, offspring probabilities

$$\begin{aligned} g_0 &= \frac{Q(1 - p_t) - (1 - p_t)}{p_t(1 - Q'(1 - p_t))}, \quad g_1 = 0, \\ g_m &= \frac{p_t^{m-1}}{m!} Q^{(m)}(1 - p_t) (1 - Q'(1 - p_t))^{-1} \quad (m \geq 2), \end{aligned} \tag{20}$$

where $p_t = p_t(\lambda, \varphi) = \mathbb{P}(\mathcal{S}_t(\varphi, T) \neq \phi)$, and generating function

$$G(z) = z + \frac{Q((1 - p_t) + p_t z) - (1 - p_t) - p_t z}{p_t(1 - Q'(1 - p_t))}. \tag{21}$$

Moreover, if $\mu(T \in \cdot)$ is critical, then so is $\nu(T \in \cdot \mid T \neq \phi)$.

An alternative proof of Lemma 2.5 can be found in Appendix C. Since Lemma 2.5 deals with the finite-leaf trees ($\text{LEAVES}(T) < \infty$), for this lemma and its proof, as well as the whole set-up of generalized dynamical pruning, we do not need to introduce the Gromov–Hausdorff metric and to require the function $\varphi : \mathcal{L} \rightarrow \mathbb{R}_+$ to be Borel measurable.

2.7. Bernoulli leaf coloring

Duquesne and Winkel considered the following type of tree reduction in [8]. Fix a probability $p \in [0, 1)$. For a finite tree $T \in \mathcal{T}^{\downarrow}$ (or \mathcal{L}^{\downarrow}), select a subset of its leaves via performing LEAVES(T) independent Bernoulli trials, where each leaf is independently selected in with probability $1 - p$. Let $\mathcal{C}_p(T)$ be the minimal subtree of T that contains all selected leaves and the root ρ . If T is a random tree, then so is $\mathcal{C}_p(T)$. Notice that \mathcal{C}_p is a random operator induced by a countable sequence of independent Bernoulli random variables.

Theorem 2.4. (Evolution of Galton–Watson trees under Bernoulli leaf coloring [8].) *Consider a critical or subcritical Galton–Watson measure $\mu \equiv \mathcal{GW}(\{q_k\})$ ($q_1 = 0$) on \mathcal{T}^{\downarrow} with generating function $Q(z)$. Then, for a given $p \in [0, 1)$, $\mu(\mathcal{C}_p(T) \in \cdot \mid \mathcal{C}_p(T) \neq \emptyset)$ is a Galton–Watson tree measure over \mathcal{T}^{\downarrow} with the generating function*

$$G_p(z) = z + \frac{Q((1 - p) + g_p z) - (1 - g_p) - g_p z}{g_p(1 - Q'(1 - g_p))}, \tag{22}$$

where $g_p = \mathbb{P}(\mathcal{C}_p(T) \neq \emptyset)$.

Theorem 2.4 readily implies that the IGW trees are invariant with respect to Bernoulli leaf coloring.

3. Results

3.1. Metric properties of invariant Galton–Watson trees

Here we derive explicit formulas for selected metric properties of $\mathcal{IGW}(q)$ and $\mathcal{IGW}(q, \lambda)$ trees in the respective spaces \mathcal{T}^{\downarrow} and \mathcal{L}^{\downarrow} . This includes the distributions of the tree height (Theorem 3.1), the tree length (Theorem 3.2), and the tree size (number of edges) (Theorem 3.3), as well as the tail asymptotics for the distributions of the tree length (Proposition 3.1) and tree size (Proposition 3.2). The proofs are collected in Section 4.1.

Theorem 3.1. (Tree height distribution.) *Let $T \in \mathcal{L}^{\downarrow}$ be an IGW tree with parameters $q \in [1/2, 1)$ and $\lambda > 0$, i.e., $T \stackrel{d}{\sim} \mathcal{IGW}(q, \lambda)$. Then the height of the tree T has the cumulative distribution function*

$$H(x) = \mathbb{P}(\text{HEIGHT}(T) \leq x) = 1 - (\lambda(1 - q)x + 1)^{-q/(1-q)}, \quad x \geq 0.$$

Notice that for the case $q = 1/2$, we have $H(x) = \frac{\lambda x}{\lambda x + 2}$, which matches the result in [20].

Theorem 3.2. (Tree length distribution.) *Let $T \in \mathcal{L}^{\downarrow}$ be an IGW tree with parameters $q \in [1/2, 1)$ and $\lambda > 0$, i.e., $T \stackrel{d}{\sim} \mathcal{IGW}(q, \lambda)$. Then the length of the tree T has the probability density function*

$$\ell(x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\Gamma(n/q + 1)}{n! (n - 1)! \Gamma(n/q - n + 2)} (\lambda q)^n x^{n-1}, \quad x \geq 0, \tag{23}$$

and the cumulative distribution function

$$L(x) = \mathbb{P}(\text{LENGTH}(T) \leq x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\Gamma(n/q + 1)}{n! n! \Gamma(n/q - n + 2)} (\lambda q)^n x^n, \quad x \geq 0. \tag{24}$$

Example 3.1. Let $q = \frac{1}{2}$. Then $\ell(x)$ is already known (see [19, 20]):

$$\ell(x) = \frac{1}{x} e^{-\lambda x} I_1(\lambda x) = \sum_{n=0}^{\infty} \frac{\lambda^{2n+1} x^{2n} e^{-\lambda x}}{2^{2n+1} (n+1)! n!}. \quad (25)$$

Next, we use the multinomial approach to show that (25) matches Equation (23) for $q = \frac{1}{2}$. First, we rewrite (25):

$$\begin{aligned} \ell(x) &= e^{-\lambda x} \sum_{n=0}^{\infty} \frac{\lambda^{2n+1}}{2^{2n+1} (n+1)! n!} x^{2n} \\ &= \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} x^k \sum_{n=0}^{\infty} \frac{\lambda^{2n+1}}{2^{2n+1} (n+1)! n!} x^{2n} \\ &= \sum_{m=0}^{\infty} \left(\sum_{k+2n=m} \frac{(-1)^k 2^{-2n-1}}{k! (n+1)! n!} \right) \lambda^{m+1} x^m \\ &= \sum_{m=0}^{\infty} \left(\sum_{k+2n=m} \frac{(-2)^k}{k! (n+1)! n!} \right) \frac{\lambda^{m+1} x^m}{2^{m+1}}. \end{aligned} \quad (26)$$

Recall that

$$(z + z^{-1} + a)^{m+1} = \sum_{n+k+j=m+1} \frac{(m+1)!}{n! k! j!} z^n z^{-j} a^k,$$

and

$$\frac{1}{2\pi i} \oint_{|z|=1} z^{n-j} dz = \delta_{j,n+1},$$

and therefore

$$\begin{aligned} &\frac{1}{2\pi i} \oint_{|z|=1} (z + z^{-1} + a)^m dz \\ &= \sum_{n+k+j=m+1} \frac{(m+1)!}{n! k! j!} a^k \frac{1}{2\pi i} \oint_{|z|=1} z^{n-j} dz \\ &= \sum_{k+2n=m} \frac{(m+1)!}{n! (n+1)! k!} a^k, \end{aligned}$$

implying

$$\sum_{k+2n=m} \frac{1}{n! (n+1)! k!} a^k = \frac{1}{2\pi i (m+1)!} \oint_{|z|=1} (z + z^{-1} + a)^{m+1} dz.$$

Now,

$$\begin{aligned} & \frac{1}{2\pi i} \oint_{|z|=1} (z + z^{-1} - 2)^{m+1} dz \\ &= \frac{1}{2\pi i} \oint_{|z|=1} \frac{(z - 1)^{2m+2}}{z^{m+1}} dz \\ &= \frac{1}{2\pi i} \oint_{|z|=1} \sum_{j=0}^{2m+2} \binom{2m+2}{j} (-1)^j z^{j-m-1} dz \\ &= (-1)^m \binom{2m+2}{m}. \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{k+2n=m} \frac{(-2)^k}{k!(n+1)!n!} &= \frac{1}{(m+1)!} \frac{1}{2\pi i} \oint_{|z|=1} (z + z^{-1} - 2)^{m+1} dz \\ &= (-1)^m \frac{1}{(m+1)!} \binom{2m+2}{m}. \end{aligned} \tag{27}$$

Thus, substituting (27) into (26), we obtain

$$\begin{aligned} \ell(x) &= \sum_{m=0}^{\infty} (-1)^m \frac{1}{(m+1)!} \binom{2m+2}{m} \frac{\lambda^{m+1} x^m}{2^{m+1}} \\ &= \sum_{m=0}^{\infty} (-1)^m \frac{(2m+2)!}{(m+1)! m! (m+2)!} (\lambda q)^{m+1} x^m \\ &= \sum_{m=0}^{\infty} (-1)^m \frac{\Gamma((m+1)/q + 1)}{(m+1)! m! \Gamma((m+1)/q - m + 1)} (\lambda q)^{m+1} x^m \end{aligned}$$

for $q = \frac{1}{2}$, as in Equation (23) from Theorem 3.2.

The following proposition is needed because computing the cumulative distribution function $L(x)$ in (24) becomes difficult (even numerically) for all values of $q \neq \frac{1}{2}$, i.e., for $q \in (1/2, 1)$.

Proposition 3.1. (Tail of the tree length distribution.) *Let $T \stackrel{d}{\sim} \mathcal{IGW}(q, \lambda)$ be an IGW tree in \mathcal{L}^1 with parameters $q \in [1/2, 1)$ and $\lambda > 0$. Then the cumulative distribution function $L(x)$ in (24) satisfies*

$$1 - L(x) \sim \frac{1}{(\lambda q)^q \Gamma(1 - q)} x^{-q}. \tag{28}$$

Example 3.2. For $q = \frac{1}{2}$, $L(x)$ is expressed as follows [19, 20]:

$$L(x) = 1 - e^{-\lambda x} (I_0(\lambda x) + I_1(\lambda x)).$$

Thus, since $I_a(z) \sim \frac{1}{\sqrt{2\pi z}} e^z$ for all $a \geq 0$, we have

$$1 - L(x) = e^{-\lambda x} (I_0(\lambda x) + I_1(\lambda x)) \sim \sqrt{\frac{2}{\lambda \pi}} x^{-1/2} = \frac{1}{(\lambda q)^q \Gamma(1 - q)} x^{-q} \quad \text{for } q = \frac{1}{2}$$

as $\Gamma(1/2) = \sqrt{\pi}$. This matches the general case in Proposition 3.1.

The following is a discrete analogue of Theorem 3.2.

Theorem 3.3. (Tree size distribution.) *Let $T \in \mathcal{T}^1$ be an IGW tree with parameters $q \in [1/2, 1)$, i.e., $T \stackrel{d}{\sim} \mathcal{IGW}(q)$. Then the number of edges in T is distributed with the probability mass function*

$$\alpha(n) = \sum_{k=1}^n (-1)^{k-1} \binom{n-1}{k-1} \frac{\Gamma(k/q+1)}{k! \Gamma(k/q-k+2)} q^k \quad \text{for } n = 1, 2, \dots, \quad (29)$$

with the cumulative distribution function

$$\mathcal{A}(x) = \sum_{k=1}^{\lfloor x \rfloor} (-1)^{k-1} \binom{\lfloor x \rfloor}{k} \frac{\Gamma(k/q+1)}{k! \Gamma(k/q-k+2)} q^k, \quad x \geq 1. \quad (30)$$

The next proposition is analogous to Proposition 3.1 and has a similar proof. It gives an estimate on the tail distribution $1 - \mathcal{A}(x)$.

Proposition 3.2. (Tail of the tree size distribution.) *Let $T \stackrel{d}{\sim} \mathcal{IGW}(q)$ be an IGW tree in \mathcal{T}^1 with parameters $q \in [1/2, 1)$. Then the cumulative distribution function $\mathcal{A}(x)$ in (30) satisfies*

$$1 - \mathcal{A}(x) \sim \frac{1}{q^q \Gamma(1 - q)} x^{-q}. \quad (31)$$

3.2. Invariance under generalized dynamical pruning

Here we consider invariance (Proposition 3.3) and uniqueness (Lemma 3.1) properties of $\mathcal{IGW}(q, \lambda)$ measures under generalized dynamical prunings. Although both Proposition 3.3 and Lemma 3.1 follow immediately from the results of Duquesne and Winkel [9, Section 3.2.1], alternative proofs of these statements that do not rely on a real tree setting are presented in Section 4.2.

We say that a Galton–Watson tree measure μ is invariant under the operation of pruning $\mathcal{S}_t(\cdot) = \mathcal{S}_t(\varphi, \cdot)$ if for $T \stackrel{d}{\sim} \mu$,

$$\mathbb{P}(\text{SHAPE}(\mathcal{S}_t(T)) = \tau \mid \mathcal{S}_t(T) \neq \emptyset) = \mu(\text{SHAPE}(T) = \tau), \quad \text{for all } \tau \in \mathcal{T}^1.$$

Proposition 3.3. (Invariance with respect to generalized dynamical pruning.) *Let $T \stackrel{d}{\sim} \mathcal{IGW}(q, \lambda)$ be an IGW tree with parameters $q \in [1/2, 1)$ and $\lambda > 0$. Then, for any monotone nondecreasing function $\varphi : \mathcal{L}^1 \rightarrow \mathbb{R}_+$ and any $t > 0$, we have*

$$T^t := \{\mathcal{S}_t(\varphi, T) \mid \mathcal{S}_t(\varphi, T) \neq \emptyset\} \stackrel{d}{\sim} \mathcal{IGW}(q, \mathcal{E}_t(\lambda)),$$

where $\mathcal{E}_t(\lambda) = \lambda p_t^{(1-q)/q}$ and $p_t = p_t(\lambda, \varphi) = \mathbb{P}(\mathcal{S}_t(\varphi, T) \neq \emptyset)$.

In other words, Proposition 3.3 yields the invariance of an $\mathcal{IGW}(q, \lambda)$ measure under generalized dynamical prunings \mathcal{S}_t . For $\varphi(T) = \text{ord}(T) - 1$, Proposition 3.3 yields the ‘if’ part of Theorem 2.1. Next, we formulate the following uniqueness result.

Lemma 3.1. (Uniqueness of IGW measures.) *Consider a critical Galton–Watson tree measure $\mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$ ($q_1 = 0$) on \mathcal{L}^1 , and let $T \stackrel{d}{\sim} \mu$. Let $\varphi : \mathcal{L} \rightarrow \mathbb{R}_+$ be a monotone nondecreasing function such that $p_t = \mathbb{P}(\mathcal{S}_t(\varphi, T) \neq \emptyset)$ is a decreasing function of t , mapping $[0, \infty)$ onto $(0, 1]$. Then μ is invariant under the operation of pruning $\mathcal{S}_t(T) = \mathcal{S}_t(\varphi, T)$ if and only if $\mu \equiv \mathcal{IGW}(q_0, \lambda)$.*

Notice that Lemma 3.1 does not imply the uniqueness result in Theorem 2.1, which is valid under the regularity assumption, Assumption 2.1. Next, we list some examples where the assumptions of Lemma 3.1 are satisfied.

Example 3.3. Let $\varphi(T) = \text{HEIGHT}(T)$. Consider a critical Galton–Watson tree measure $\mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$ ($q_1 = 0$) on \mathcal{L}^1 , and let $T \stackrel{d}{\sim} \mu$. Then $1 - p_t = P_{1,0}(t)$ is the probability of extinction by time t of the critical continuous-time branching process. Since $P_{1,0}(t)$ is a continuous function of t that maps $[0, \infty)$ onto $[0, 1)$, Lemma 3.1 implies that $\mathcal{IGW}(q, \lambda)$ is the only class of Galton–Watson measures that are invariant under the generalized dynamical pruning with $\varphi(T) = \text{HEIGHT}(T)$.

Example 3.4. Let $\varphi(T) = \text{LENGTH}(T)$. Consider a critical Galton–Watson tree measure $\mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$ ($q_1 = 0$) on \mathcal{L}^1 , and let $T \stackrel{d}{\sim} \mu$. Denote by N the number of edges in T . Then the density function of $\text{LENGTH}(T)$ can be expressed as $\sum_{k=1}^{\infty} \mathbb{P}(N = k) f_{k,\lambda}(x)$, where $f_{k,\lambda}(x)$ is a gamma function $f_{k,\lambda}(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x}$. Hence, the cumulative distribution function of $\text{LENGTH}(T)$,

$$\mathbb{P}(\text{LENGTH}(T) \leq t) = 1 - p_t,$$

is a continuous function of t mapping $[0, \infty)$ onto $[0, 1)$. Thus, by Lemma 3.1, $\mathcal{IGW}(q, \lambda)$ is the only class of Galton–Watson measures invariant under the generalized dynamical pruning with $\varphi(T) = \text{LENGTH}(T)$.

Next, we check that Proposition 3.3 and Theorem 3.1 are consistent with this semigroup property of the generalized dynamical pruning induced by $\varphi(T) = \text{HEIGHT}(T)$ as in Example 2.2. Indeed, for $T \stackrel{d}{\sim} \mathcal{IGW}(q, \lambda)$, Proposition 3.3 yields

$$T^t := \{\mathcal{S}_t(\varphi, T) \mid \mathcal{S}_t(\varphi, T) \neq \emptyset\} \stackrel{d}{\sim} \mathcal{IGW}(q, \mathcal{E}_t(\lambda)),$$

where by Theorem 3.1, $\mathcal{E}_t(\lambda) = \lambda p_t^{(1-q)/q} = \frac{\lambda}{\lambda(1-q)t+1}$. Hence,

$$\mathcal{E}_s \circ \mathcal{E}_t(\lambda) = \mathcal{E}_s(\mathcal{E}_t(\lambda)) = \mathcal{E}_{t+s}(\lambda),$$

which reaffirms the semigroup property of \mathcal{S}_t for $\varphi(T) = \text{HEIGHT}(T)$.

3.3. Invariant Galton–Watson trees $\mathcal{IGW}(q)$ as attractors

The following result extends Theorem 2.2 to all generalized dynamical pruning operators $\mathcal{S}_t(T) = \mathcal{S}_t(\varphi, T)$.

Theorem 3.4. (IGW attractors under generalized dynamical pruning.) *Consider a Galton–Watson measure $\mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$ with $q_1 = 0$ on \mathcal{L}^1 . Suppose the measure is critical and*

Assumption 2.1 is satisfied. Then, for any random tree $T \in \mathcal{L}^1$ distributed according to μ , i.e., $T \stackrel{d}{\sim} \mu$,

$$\lim_{t \rightarrow \infty} \mathbb{P}(\text{SHAPE}(\mathcal{S}_t(T)) = \tau \mid \mathcal{S}_t(T) \neq \phi) = \mu^*(\tau), \quad \text{for all } \tau \in \mathcal{T}^1,$$

where μ^* denotes the IGW measure $\mathcal{IGW}(q)$ with $q = \frac{1}{2-L}$ and L defined in (7).

Finally, suppose the Galton–Watson measure $\mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$ (with $q_1 = 0$) is subcritical; then for $T \stackrel{d}{\sim} \mu$, the distribution $\mathbb{P}(\text{SHAPE}(\mathcal{S}_t(T)) = \cdot \mid \mathcal{S}_t(T) \neq \phi)$ converges to a point mass measure on the tree reduced to a stem, $\mathcal{GW}(q_0=1)$.

Theorem 3.4 is proved in Section 4.3.

The next two corollaries of Theorem 3.4 follow immediately from Lemmas 2.2 and 2.4.

Corollary 3.1. (Attraction property of critical Galton–Watson trees of Zipf type.) Consider a critical Galton–Watson process $\mu \equiv \mathcal{GW}(\{q_k\})$ with $q_1 = 0$, with offspring distribution q_k of Zipf type, i.e., $q_k \sim Ck^{-(\alpha+1)}$, with $\alpha \in (1, 2]$ and $C > 0$. Then, for a random tree $T \in \mathcal{L}^1$ distributed according to μ , i.e., $T \stackrel{d}{\sim} \mu$,

$$\lim_{t \rightarrow \infty} \mathbb{P}(\text{SHAPE}(\mathcal{S}_t(T)) = \tau \mid \mathcal{S}_t(T) \neq \phi) = \mu^*(\tau), \quad \text{for all } \tau \in \mathcal{T}^1,$$

where μ^* is the IGW measure $\mathcal{IGW}\left(\frac{1}{\alpha}\right)$.

Corollary 3.2. (Attraction property of critical binary Galton–Watson tree [5].) Consider a critical Galton–Watson process $\mu \equiv \mathcal{GW}(\{q_k\})$ with $q_1 = 0$. Assume one of the following two conditions holds:

(a) The second moment assumption is satisfied:

$$\sum_{k=2}^{\infty} k^2 q_k < \infty.$$

(b) Assumption 2.1 is satisfied, and the ‘2–’ moment assumption is satisfied, i.e.,

$$\sum_{k=2}^{\infty} k^{2-\epsilon} q_k < \infty \quad \forall \epsilon > 0.$$

Then, for a random tree $T \in \mathcal{L}^1$ distributed according to μ , i.e., $T \stackrel{d}{\sim} \mu$,

$$\lim_{t \rightarrow \infty} \mathbb{P}(\text{SHAPE}(\mathcal{S}_t(T)) = \tau \mid \mathcal{S}_t(T) \neq \phi) = \mu^*(\tau), \quad \text{for all } \tau \in \mathcal{T}^1,$$

where μ^* is the critical binary Galton–Watson measure $\mathcal{IGW}(1/2)$.

Next, we state a result for the Bernoulli leaf coloring operator \mathcal{C}_p (see Section 2.7), which is analogous to the one in Theorem 3.4.

Theorem 3.5. (IGW attractors under Bernoulli leaf coloring.) Consider a Galton–Watson measure $\mu \equiv \mathcal{GW}(\{q_k\})$ with $q_1 = 0$ on \mathcal{T}^1 . Suppose the measure is critical and Assumption 2.1 is satisfied. Then, for a random tree $T \in \mathcal{T}^1$ distributed according to μ , i.e., $T \stackrel{d}{\sim} \mu$,

$$\lim_{p \rightarrow 1^-} \mathbb{P}(\mathcal{C}_p(T) = \tau \mid \mathcal{C}_p(T) \neq \phi) = \mu^*(\tau), \quad \text{for all } \tau \in \mathcal{T}^1,$$

where μ^* denotes the IGW measure $\mathcal{IGW}(q)$ with $q = \frac{1}{2-L}$ and L as defined in (7).

Suppose $\mu \equiv \mathcal{GW}(\{q_k\})$ (with $q_1 = 0$) is subcritical; then for $T \stackrel{d}{\sim} \mu$, the conditional distribution $\mathbb{P}(\mathcal{C}_p(T) = \cdot \mid \mathcal{C}_p(T) \neq \phi)$ converges to a point mass measure on the tree reduced to the stem, $\mathcal{GW}(q_0=1)$.

Theorem 3.5 is proved in Section 4.3.

Finally, another result analogous to Theorem 3.4 can be obtained for iterative hereditary reductions (see Section 2.6).

Theorem 3.6. (IGW attractors under generalized hereditary reductions.) *Consider a Galton–Watson measure $\mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$ with $q_1 = 0$ on \mathcal{L}^1 . Suppose the measure is critical and Assumption 2.1 is satisfied. Let $T \in \mathcal{L}^1$ be a random tree distributed according to μ , and let H_1, H_2, \dots be a sequence of hereditary properties satisfying*

$$\lim_{n \rightarrow \infty} \mathbb{P}(R_{H_n} \circ \dots \circ R_{H_1}(T) \neq \phi) = 0,$$

where R_{H_1}, R_{H_2}, \dots are the corresponding hereditary reductions. Then, for $T \stackrel{d}{\sim} \mu$,

$$\lim_{t \rightarrow \infty} \mathbb{P}(\text{SHAPE}(\mathcal{S}_t(T)) = \tau \mid \mathcal{S}_t(T) \neq \phi) = \mu^*(\tau), \quad \text{for all } \tau \in \mathcal{T}^1,$$

where μ^* denotes the IGW measure $\mathcal{IGW}(q)$ with $q = \frac{1}{2-L}$ and L as defined in (7).

If μ is a subcritical Galton–Watson measure, then for $T \stackrel{d}{\sim} \mu$, the conditional distribution $\mathbb{P}(\text{SHAPE}(\mathcal{S}_t(T)) = \cdot \mid \mathcal{S}_t(T) \neq \phi)$ converges to a point mass measure on the tree reduced to the stem, $\mathcal{GW}(q_0=1)$.

Theorem 3.6 is proved in Section 4.3.

4. Proofs

4.1. Metric properties of invariant Galton–Watson trees

Proof of Theorem 3.1. Consider a tree $T \stackrel{d}{\sim} \mathcal{IGW}(q, \lambda)$. Let X denote the length of the stem connecting the random tree’s root ρ to the root’s only child vertex v_0 . Let $K = \text{br}(v_0)$ be the branching number of v_0 , and let the K subtrees branching out of v_0 be denoted by $T_i, 1 \leq i \leq K$. Let $H(x)$ be the cumulative distribution function for the height of T . Then, for each subtree $T_i \stackrel{d}{\sim} \mathcal{IGW}(q, \lambda)$, its height $\text{HEIGHT}(T_i)$ has the same cumulative distribution function $H(x)$. The number of subtrees $K \stackrel{d}{\sim} q_k$ has generating function $Q(z) = z + q(1 - z)^{1/q}$. Let $M(x)$ denote the cumulative distribution function of $\max_{1 \leq i \leq K} \{\text{HEIGHT}(T_i)\}$; then

$$\begin{aligned} M(x) &= \mathbb{P}\left(\max_{1 \leq i \leq K} \{\text{HEIGHT}(T_i)\} \leq x\right) = \sum_{k=0}^{\infty} q_k \mathbb{P}\left(\max_{1 \leq i \leq K} \{\text{HEIGHT}(T_i)\} \leq x \mid K = k\right) \\ &= \sum_{k=0}^{\infty} q_k \mathbb{P}(\text{HEIGHT}(T) \leq x)^k = \sum_{k=0}^{\infty} q_k (H(x))^k \\ &= (Q \circ H)(x) = H(x) + q(1 - H(x))^{1/q}. \end{aligned} \tag{32}$$

The stem length X is an exponentially distributed random variable with parameter λ and density function $\varphi_\lambda(x) = \lambda \exp\{-\lambda x\} \mathbf{1}_{x \geq 0}$. Since $\text{HEIGHT}(T) = X + \max_{1 \leq i \leq K} \{\text{HEIGHT}(T_i)\}$, we have

$$H(x) = \varphi_\lambda * M(x). \quad (33)$$

We will use the following notation: let $\widehat{g}(t) = \int_{-\infty}^{\infty} e^{itx} g(x) dx$ denote the Fourier transform of $g(x)$. Equations (32) and (33) yield

$$H(x) = \varphi_\lambda * (Q \circ H)(x).$$

Taking the Fourier transform, we obtain

$$\widehat{H}(t) = \frac{\lambda}{\lambda - it} \left(\widehat{H}(t) + q \widehat{(1 - H)^{1/q}}(t) \right),$$

which simplifies to

$$it\widehat{H}(t) + \lambda q \widehat{(1 - H)^{1/q}}(t) = 0,$$

where

$$\widehat{(1 - H)^{1/q}}(t) = \int_{-\infty}^{\infty} e^{itx} (1 - H(x))^{1/q} dx.$$

Therefore,

$$\int_{-\infty}^{\infty} e^{itx} \left(itH(x) + \lambda q (1 - H(x))^{1/q} \right) dx = 0 \quad \forall t \in \mathbb{R}, \quad (34)$$

where integration by parts yields

$$\int_{-\infty}^{\infty} e^{itx} itH(x) dx = - \int_{-\infty}^{\infty} e^{itx} H'(x) dx. \quad (35)$$

Substituting (35) back into (34) yields

$$\int_{-\infty}^{\infty} e^{itx} \left(H'(x) - \lambda q (1 - H(x))^{1/q} \right) dx = 0 \quad \forall t \in \mathbb{R},$$

which by Parseval's equation implies the following ordinary differential equation:

$$H'(x) = \lambda q (1 - H(x))^{1/q}. \quad (36)$$

Next, we solve the differential equation (36) above via integration, obtaining

$$H(x) = 1 - ((\lambda x + C)(1 - q))^{-\frac{q}{1-q}}, \quad (37)$$

where C is a scalar. Since $H(x)$ is the cumulative distribution function of a positive random variable $\text{HEIGHT}(T)$, we have $H(0) = 0$, implying $C = \frac{1}{1-q}$. Thus, for $q \in [\frac{1}{2}, 1)$,

$$H(x) = 1 - \left(\left(\lambda x + \frac{1}{1-q} \right) (1-q) \right)^{-\frac{q}{1-q}} = 1 - (\lambda(1-q)x + 1)^{-q/(1-q)}. \quad \square$$

Next, we use the following application of the Lagrange inversion theorem (Theorem A.1).

Lemma 4.1. *Let $q \in [1/2, 1)$ be given. Suppose $W = W(z)$ is an analytic function satisfying the equation*

$$z = \frac{W}{(1-W)^{1/q}}$$

in a neighborhood of the origin, where we take the $-\pi < \arg(z) < \pi$ branch of the function $z^{1/q}$. Then, for z near the origin, we have

$$W = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\Gamma(n/q + 1)}{n! \Gamma(n/q - n + 2)} z^n. \quad (38)$$

Observe that the conclusion of Lemma 4.1 also applies in a real-valued setting, under the assumption of infinite differentiability of $W : \mathbb{R} \rightarrow \mathbb{R}$. Here, if $z = \frac{W}{(1-W)^{1/q}}$ for $z \in \mathbb{R}$ in a neighborhood of the origin on the real line, then the power series expansion (38) holds in proximity to 0.

Proof of Lemma 4.1. We notice that the function $f(w) = \frac{w}{(1-w)^{1/q}}$ is analytic at $w = 0$, and $f'(0) = 1 \neq 0$. Thus, we can apply the Lagrange inversion theorem (Theorem A.1) to express W in terms of z power series. Now, since

$$\left(\frac{w}{f(w)} \right)^n = (1-w)^{n/q},$$

we have

$$\begin{aligned} \frac{d^{n-1}}{dw^{n-1}} \left(\frac{w}{f(w)} \right)^n \Big|_{w=0} &= (-1)^{n-1} (n/q)(n/q-1) \dots (n/q-n+2) \\ &= (-1)^{n-1} \frac{\Gamma(n/q + 1)}{\Gamma(n/q - n + 2)}. \end{aligned}$$

Therefore, by the Lagrange inversion theorem (Theorem A.1), we obtain

$$W = \sum_{n=1}^{\infty} \frac{z^n}{n!} \left[\frac{d^{n-1}}{dw^{n-1}} \left(\frac{w}{f(w)} \right)^n \right]_{w=0} = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\Gamma(n/q + 1)}{n! \Gamma(n/q - n + 2)} z^n. \quad \square$$

Proof of Theorem 3.2. Consider a tree $T \stackrel{d}{\sim} \mathcal{IGW}(q, \lambda)$ consisting of a stem of length X that connects the root ρ to its child vertex v_0 , and $K = \text{br}(v_0)$ subtrees T_i , $1 \leq i \leq K$, branching out from v_0 . Let $\ell(x)$ be the density function of the length of T . Notice that the length of each subtree T_i is also $\ell(x)$ -distributed. The random variable $K \stackrel{d}{\sim} q_k$ has generating function $Q(z) =$

$z + q(1 - z)^{1/q}$. Letting $N(x)$ denote the probability density function of $\sum_{1 \leq i \leq K} \{\text{LENGTH}(T_i)\}$, we have

$$N(x) = \sum_{k=0}^{\infty} q_k \ell_k(x), \quad \text{where } \ell_k(x) = \underbrace{\ell * \dots * \ell}_{k \text{ times}}(x). \quad (39)$$

Observe that $\text{LENGTH}(T) = X + \sum_{1 \leq i \leq K} \{\text{LENGTH}(T_i)\}$, where X has exponential probability density function $\varphi_\lambda(x) = \lambda \exp\{-\lambda x\} \mathbf{1}_{x \geq 0}$. Thus, $\ell(x)$ can be represented as the following convolution:

$$\ell(x) = \varphi_\lambda * N(x). \quad (40)$$

For $t \geq 0$, let the function $\mathcal{L}[g](t) = \int_0^\infty e^{-tx} g(x) dx$ denote the Laplace transform of g . Then (39) and (40) imply

$$\mathcal{L}[\ell](t) = \mathcal{L}[\varphi_\lambda](t) \mathcal{L}[N](t) = \mathcal{L}[\varphi_\lambda](t) \mathcal{Q}(\mathcal{L}[\ell](t)) = \frac{\lambda}{\lambda + t} \left(\mathcal{L}[\ell](t) + q(1 - \mathcal{L}[\ell](t))^{1/q} \right),$$

which simplifies to

$$t\mathcal{L}[\ell](t) = \lambda q(1 - \mathcal{L}[\ell](t))^{1/q}. \quad (41)$$

Letting $z = \frac{\lambda q}{t}$ and $\Lambda = \mathcal{L}[\ell] \left(\frac{\lambda q}{z} \right) = \mathcal{L}[\ell](t)$, we have

$$z = \frac{\Lambda}{(1 - \Lambda)^{1/q}}.$$

Then Lemma 4.1 yields

$$\mathcal{L}[\ell](t) = \Lambda = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\Gamma(n/q + 1)}{n! \Gamma(n/q - n + 2)} \frac{(\lambda q)^n}{t^n}.$$

Finally, we invert the Laplace transform $\mathcal{L}[\ell](t)$, obtaining

$$\ell(x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\Gamma(n/q + 1)}{n! (n-1)! \Gamma(n/q - n + 2)} (\lambda q)^n x^{n-1}. \quad \square$$

Proof of Proposition 3.1. Observe that

$$\begin{aligned} 1 - \mathcal{L}[\ell](t) &= \int_0^\infty (1 - e^{-tx}) \ell(x) dx = t \int_0^\infty \int_0^x e^{-ty} \ell(x) dy dx \\ &= t \int_0^\infty e^{-ty} \int_y^\infty \ell(x) dx dy = t \int_0^\infty e^{-ty} (1 - L(y)) dy = t \mathcal{L}[1-L](t). \end{aligned}$$

Thus, by (41), we have

$$t\mathcal{L}[\ell](t) = \lambda q(1 - \mathcal{L}[\ell](t))^{1/q} = \lambda q t^{1/q} (\mathcal{L}[1-L](t))^{1/q},$$

and therefore,

$$\mathcal{L}[1-L](t) = \frac{1}{t^{1-q}} \frac{(\mathcal{L}[\ell](t))^q}{(\lambda q)^q}, \quad \text{where} \quad \lim_{t \rightarrow 0^+} \frac{(\mathcal{L}[\ell](t))^q}{(\lambda q)^q} = \frac{1}{(\lambda q)^q}.$$

Hence, by the Hardy–Littlewood–Karamata Tauberian theorem for Laplace transforms [12],

$$1 - L(x) \sim \frac{1}{(\lambda q)^q \Gamma(1 - q)} x^{-q}. \quad \square$$

Proof of Theorem 3.3. Observe that in a reduced tree $T \in \mathcal{T}^1 \setminus \{\phi\}$, the number of edges equals one plus the number of edges in all subtrees splitting from the stem. Therefore, $\alpha(0) = 0$, $\alpha(1) = q_0$, and

$$\alpha(n + 1) = \sum_{k=1}^n q_k \underbrace{\alpha * \dots * \alpha}_{k \text{ times}}(n), \quad n = 1, 2, \dots$$

Therefore, the generating function $a(z) = \sum_{n=1}^{\infty} z^n \alpha(n)$ satisfies $a(z) = zQ(a(z))$. Hence,

$$a(z) = z \left(a(z) + q(1 - a(z))^{1/q} \right), \tag{42}$$

and therefore,

$$w = \frac{a}{(1 - a)^{1/q}}, \quad \text{where } a = a(z) \text{ and } w = \frac{qz}{1 - z}.$$

Lemma 4.1 yields

$$\begin{aligned} a(z) &= \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} (-1)^{k-1} \binom{n-1}{k-1} \frac{\Gamma(k/q + 1)}{k! \Gamma(k/q - k + 2)} q^k z^n \\ &= \sum_{n=1}^{\infty} z^n \sum_{k=1}^n (-1)^{k-1} \binom{n-1}{k-1} \frac{\Gamma(k/q + 1)}{k! \Gamma(k/q - k + 2)} q^k. \end{aligned}$$

Thus, since $a(z) = \sum_{n=1}^{\infty} z^n \alpha(n)$, Equation (29) follows. Finally, the cumulative distribution function equals

$$\begin{aligned} \mathcal{A}(x) &= \sum_{n=1}^{\lfloor x \rfloor} \alpha(n) = \sum_{n=1}^{\lfloor x \rfloor} \sum_{k=1}^n (-1)^{k-1} \binom{n-1}{k-1} \frac{\Gamma(k/q + 1)}{k! \Gamma(k/q - k + 2)} q^k \\ &= \sum_{k=1}^{\lfloor x \rfloor} (-1)^{k-1} \left(\sum_{n=k}^{\lfloor x \rfloor} \binom{n-1}{k-1} \right) \frac{\Gamma(k/q + 1)}{k! \Gamma(k/q - k + 2)} q^k \\ &= \sum_{k=1}^{\lfloor x \rfloor} (-1)^{k-1} \binom{\lfloor x \rfloor}{k} \frac{\Gamma(k/q + 1)}{k! \Gamma(k/q - k + 2)} q^k \end{aligned}$$

for all real $x \geq 1$, and (30) holds. □

4.2. Invariance under generalized dynamical pruning

Proof of Proposition 3.3. For $q \in [1/2, 1)$ and $Q(z) = z + q(1-z)^{1/q}$, Equation (21) in Lemma 2.5 implies

$$\begin{aligned} G(z) &= z + \frac{Q(1-p_t+p_tz) - (1-p_t) - zp_t}{p_t(1-Q'(1-p_t))} \\ &= z + p_t^{-1/q} (Q(z + (1-z)(1-p_t)) - (1-p_t) - zp_t) \\ &= z + p_t^{-1/q} q p_t^{1/q} (1-z)^{1/q} = Q(z). \end{aligned}$$

The rest of the proof follows from Lemma 2.5 as

$$\lambda(1-Q'(1-p_t)) = \lambda p_t^{(1-q)/q}, \quad (43)$$

yielding $S_t(T) \stackrel{d}{\sim} \mathcal{IGW}(q, \lambda p_t^{(1-q)/q})$. □

Proof of Lemma 3.1. From Lemma 2.5, we have

$$g_0 = \frac{Q(1-p_t) - (1-p_t)}{p_t(1-Q'(1-p_t))} \quad \text{and} \quad G(z) = z + \frac{Q(1-p_t+p_tz) - (1-p_t) - p_tz}{p_t(1-Q'(1-p_t))}.$$

Combining the above yields

$$G(z) = z + g_0 \frac{Q(1-p_t+p_tz) - (1-p_t) - p_tz}{Q(1-p_t) - (1-p_t)}.$$

Suppose μ is invariant under the operation of pruning $S_t(T) = S_t(\varphi, T)$; then $G(z) = Q(z)$ and $g_0 = q_0$, implying

$$Q(z) = z + q_0 \frac{Q(1-p_t+p_tz) - (1-p_t) - p_tz}{Q(1-p_t) - (1-p_t)}. \quad (44)$$

Let $R(z) = \frac{Q(z)-z}{q_0}$; then Equation (44) can be rewritten as

$$R(z) = \frac{R(1-p_t+p_tz)}{R(1-p_t)}.$$

Thus, for $\ell(z) = \ln(R(1-z))$, we have $\ell(1-z) + \ell(p_t) = \ell(p_t(1-z))$ as $1-p_t+p_tz = 1-p_t(1-z)$.

Therefore, $\ell(p_t x) = \ell(x) + \ell(p_t)$. Let $r(y) = \ell(e^y)$; then

$$r(y + \varepsilon_t) = r(y) + r(\varepsilon_t) \quad \forall t \geq 0, \quad (45)$$

where $\varepsilon_t = \ln p_t$. Here $r(0) = \ln R(0) = 0$.

We notice that the domain of $r(y)$ is $y \in (-\infty, 0]$, and

$$\{\varepsilon_t : t \in [0, \infty)\} = (-\infty, 0],$$

as $1-p_t$ is an increasing function of t mapping $[0, \infty)$ onto $[0, 1)$. Hence, Equation (45) implies the following Cauchy functional equation:

$$r(y + \varepsilon) = r(y) + r(\varepsilon) \quad \forall y, \varepsilon \in (-\infty, 0]. \quad (46)$$

The general Cauchy functional equation states that, assuming

- continuity ($f(x) \in C(\mathbb{R})$) and
- additivity ($f(x + y) = f(x) + f(y)$ for all $x, y \in \mathbb{R}$),

the function $f(x) = cx$ for some $c \in \mathbb{R}$. Notice that (46) is a sub-case of the general Cauchy functional equation restricted to a half-line, and therefore has the same linear solution and the same proof. Thus, (46) yields that $r(y) = \kappa y$ for some constant κ .

Thus, we have $\ell(x) = \kappa \ln(x)$,

$$\kappa \ln(1 - z) = \ell(1 - z) = \ln R(z) = \ln \left(\frac{Q(z) - z}{q_0} \right),$$

and

$$Q(z) = z + q_0(1 - z)^\kappa.$$

Finally, $q_1 = 0$ yields $Q'(0) = 0$. Therefore, $Q'(z) = 1 - q_0\kappa(1 - z)^{\kappa-1}$ implies $\kappa = \frac{1}{q_0}$. \square

4.3. Invariant Galton–Watson trees $\mathcal{IGW}(q)$ as attractors

First we prove the following result, related to Lemma 2.1.

Lemma 4.2. *Consider a critical Galton–Watson measure $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$. If Assumption 2.1 is satisfied, then for $g(x)$ defined in (6), the limit*

$$\lim_{x \rightarrow 1^-} \frac{(1 - x)g'(x)}{g(x)} \tag{47}$$

exists and is equal to the limit L defined in (7).

Proof. Note that for $x \in (-1, 1)$,

$$\frac{Q(x) - x}{(1 - x)(1 - Q'(x))} = \frac{1}{2 - \frac{(1-x)g'(x)}{g(x)}}.$$

Thus, by Assumption 2.1, the limit

$$\lim_{x \rightarrow 1^-} \frac{(1 - x)g'(x)}{g(x)}$$

either exists or is equal to $\pm\infty$. Hence, by L'Hôpital's rule,

$$L = \lim_{x \rightarrow 1^-} \left(\frac{\ln g(x)}{-\ln(1 - x)} \right) = \lim_{x \rightarrow 1^-} \frac{(1 - x)g'(x)}{g(x)}.$$

\square

Before proving Theorem 3.4, we will need the following result.

Lemma 4.3. *Consider a critical Galton–Watson measure $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$, let $g(x)$ be as defined in (6), and let L be as defined in (7). If Assumption 2.1 is satisfied, then $g(1 - 1/y)$ is a regularly varying function (Definition B.1) with index L , i.e.,*

$$\lim_{x \rightarrow 1^-} \frac{g\left(\left(1 - \frac{1}{r}\right) + \frac{1}{rx}\right)}{g(x)} = \lim_{y \rightarrow \infty} \frac{g\left(1 - \frac{1}{ry}\right)}{g\left(1 - \frac{1}{y}\right)} = r^L \quad \text{for all } r > 0. \tag{48}$$

Proof. For $\alpha > -L - 1$, L'Hôpital's rule and Lemma 4.2 yield

$$\begin{aligned} \lim_{y \rightarrow \infty} \frac{y^{\alpha+1} g(1-1/y)}{\int_a^y s^\alpha g(1-1/s) ds} &= \lim_{y \rightarrow \infty} \frac{(\alpha+1)y^\alpha g(1-1/y) + y^{\alpha-1} g'(1-1/y)}{y^\alpha g(1-1/y)} \\ &= \alpha + 1 + \lim_{y \rightarrow \infty} \frac{y^{\alpha-1} g'(1-1/y)}{y^\alpha g(1-1/y)} \\ &= \alpha + 1 + \lim_{x \rightarrow 1-} \frac{(1-x)g'(x)}{g(x)} = \alpha + 1 + L. \end{aligned}$$

Hence, by the converse of Karamata's theorem (Theorem B.2), $g(1-1/y)$ is a regularly varying function with index L , and (48) holds. \square

The following lemma will be the instrument for establishing that $\mathcal{IGW}(q)$ trees are attractors.

Lemma 4.4. Consider a Galton–Watson measure $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$ on \mathcal{T}^1 . Suppose the measure is critical and Assumption 2.1 is satisfied. Then its progeny generating function $Q(z)$ satisfies

$$\lim_{x \rightarrow 1-} \frac{Q(z + (1-z)x) - (z + (1-z)x)}{(1-x)(1-Q'(x))} = \frac{1}{2-L} (1-z)^{2-L},$$

where L is as defined in (7).

If the Galton–Watson measure $\mathcal{GW}(\{q_k\})$ (with $q_1 = 0$) is subcritical, then

$$\lim_{x \rightarrow 1-} \frac{Q(z + (1-z)x) - (z + (1-z)x)}{(1-x)(1-Q'(x))} = 1 - z.$$

Proof. Consider a critical Galton–Watson measure $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$ and progeny generating function $Q(z)$. For $x, z \in (-1, 1)$, we have

$$Q(z + (1-z)x) - (z + (1-z)x) = (1-z)^2 (1-x)^2 g(z + (1-z)x).$$

Thus, as

$$1 - Q'(x) = 2(1-x)g(x) - (1-x)^2 g'(x),$$

Lemma 4.2 yields

$$\begin{aligned} \lim_{x \rightarrow 1-} \frac{Q(z + (1-z)x) - (z + (1-z)x)}{(1-x)(1-Q'(x))} &= (1-z)^2 \lim_{x \rightarrow 1-} \frac{g(z + (1-z)x)}{2g(x) - (1-x)g'(x)} \\ &= (1-z)^2 \lim_{x \rightarrow 1-} \frac{g(z + (1-z)x)}{\left(2 - \frac{(1-x)g'(x)}{g(x)}\right) g(x)} = \frac{1}{2-L} (1-z)^2 \lim_{x \rightarrow 1-} \frac{g(z + (1-z)x)}{g(x)} \\ &= \frac{1}{2-L} (1-z)^2 (1-z)^{-L} = \frac{1}{2-L} (1-z)^{2-L} \end{aligned}$$

by (48) with $r = \frac{1}{1-z}$. The main statement in Lemma 4.4 follows.

Now, if we consider a subcritical Galton–Watson measure $\mathcal{GW}(\{q_k\})$ with $q_1 = 0$ and progeny generating function $Q(z)$, then $Q'(1) < 1$, and

$$\begin{aligned} & \lim_{x \rightarrow 1^-} \frac{Q(z + (1 - z)x) - (z + (1 - z)x)}{(1 - x)(1 - Q'(x))} \\ &= \frac{1}{1 - Q'(1)} \lim_{x \rightarrow 1^-} \frac{Q(z + (1 - z)x) - (z + (1 - z)x)}{1 - x} \\ &= \frac{1 - z}{1 - Q'(1)} \lim_{x \rightarrow 1^-} \frac{Q(1 - (1 - z)(1 - x)) - Q(1) + (1 - z)(1 - x)}{(1 - z)(1 - x)} = 1 - z. \end{aligned}$$

□

We are now ready to prove Theorem 3.4.

Proof of Theorem 3.4. Suppose $\mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$ with $q_1 = 0$ is critical and Assumption 2.1 holds. Then, by Equation (21) in Lemma 2.5 and Lemma 4.4, the generating function of $\text{SHAPE}(\mathcal{S}_t(T))$ converges to

$$\begin{aligned} & z + \lim_{t \rightarrow \infty} \frac{Q(z + (1 - z)(1 - p_t)) - (1 - p_t) - zp_t}{p_t(1 - Q'(1 - p_t))} \\ &= z + \lim_{x \rightarrow 1^-} \frac{Q(z + (1 - z)x) - (z + (1 - z)x)}{(1 - x)(1 - Q'(x))} \\ &= z + \frac{1}{2 - L}(1 - z)^{2-L}, \end{aligned}$$

the generating function of $\mathcal{IGW}(q)$ with $q = \frac{1}{2-L}$ and L as defined in (7).

If the Galton–Watson measure $\mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$ (with $q_1 = 0$) is subcritical, then Lemma 2.5 and Lemma 4.4 yield convergence of the generating function of $\text{SHAPE}(\mathcal{S}_t(T))$ to

$$\begin{aligned} & z + \lim_{t \rightarrow \infty} \frac{Q(z + (1 - z)(1 - p_t)) - (1 - p_t) - zp_t}{p_t(1 - Q'(1 - p_t))} \\ &= z + \lim_{x \rightarrow 1^-} \frac{Q(z + (1 - z)x) - (z + (1 - z)x)}{(1 - x)(1 - Q'(x))} \\ &= z + (1 - z) = 1, \end{aligned}$$

the generating function of $\mathcal{GW}(q_0=1)$. Hence, for $T \stackrel{d}{\sim} \mu$, the conditional distribution $\mathbb{P}(\text{SHAPE}(\mathcal{S}_t(T)) = \cdot \mid \mathcal{S}_t(T) \neq \emptyset)$ converges to a point mass measure, $\mathcal{GW}(q_0=1)$. □

Proof of Theorem 3.5. Analogously to the proof of Theorem 3.4 above, Theorem 3.5 follows from the formula (22) in Theorem 2.4 and Lemma 4.4. □

Proof of Theorem 3.6. Following the steps in the proof of Theorem 3.4 above, Theorem 3.6 follows from the formula (18) in Theorem 2.3 with $p = \mathbb{P}(R_{H_n} \circ \dots \circ R_{H_1}(T) \neq \emptyset)$ and Lemma 4.4. □

5. Discussion

In this work, we established the metric and attractor properties of the IGW branching processes with respect to a family of the generalized dynamical pruning operators. Informally,

these operators eliminate a tree from leaves toward the root and are flexible enough to accommodate a number of classic (e.g. continuous erasure) and custom (e.g. erasure by the number of leaves) tree elimination rules. Together with the richness of the IGW family, which includes power-law offspring distributions with tail indices in the range between 1 and 2, this makes the results presented a useful tool for a variety of physical and mathematical problems.

Observe that erasing a random tree in accordance with the generalized dynamical pruning describes a coalescence dynamics—the merging of particles represented by the tree leaves into consecutively larger clusters represented by the internal vertices. The invariance and attractor properties of the IGW branching processes can be used to study a number of merger dynamics. For example, the continuum ballistic annihilation process (a ballistic motion of random-velocity particles that annihilate at contact) has been shown in [19] to correspond to a generalized dynamical pruning with $\varphi(T) = \text{LENGTH}(T)$, as in Example 2.3. The invariance of the critical binary Galton–Watson measure $\mathcal{IGW}(1/2, \lambda)$ under the generalized dynamical pruning was used in [19] to obtain an explicit analytical description of the annihilation dynamics for a special case of the initial two-valued velocity alternating at the instances of a Poisson process. Similarly, the generalized dynamical pruning with $\varphi(T) = \text{HEIGHT}(T)$ as in Example 2.2 corresponds to the one-dimensional Zeldovich model in cosmology. The invariance results in this work may provide an interesting analytical insight into the dynamics of these and other models of coalescence.

The IGW branching processes arise naturally in seismological data that are traditionally modeled by branching processes with immigration; see [22] for a review and discussion. In essence, a sequence of earthquakes in a region is represented as a collection of clusters, each of which is initiated by an immigrant (the first cluster event). It has been shown in [22] that the IGW process provides a close approximation to the existing earthquake occurrence models and to the observed earthquake cluster statistics in southern California. The metric properties of the IGW trees have a meaningful interpretation in the seismological setting, where the edge lengths represent inter-event times. The attraction property of the IGW processes allows one to construct a robust earthquake modeling framework, which is stable with respect to various catalogue uncertainties. The IGW processes may provide a useful model in other areas that deal with imprecisely observed data represented by trees.

We conclude with an open problem. Lemma 3.1 established the uniqueness of the IGW processes as the invariants of the generalized dynamical pruning. The results of Duquesne and Winkel [8] (see Theorem 2.4 of this paper) show that the IGW processes are invariant with respect to a broader set of tree reductions. It would be interesting to describe all tree transforming operators that preserve the IGW invariance (and attraction) property.

Appendix A. Lagrange inversion theorem

The Lagrange inversion theorem (also known as the Lagrange inversion formula) can be found in Whittaker and Watson [31] and in Abramowitz and Stegun [2].

Theorem A.1. (Lagrange inversion theorem.) *Consider a function $g : \mathbb{C} \rightarrow \mathbb{C}$ such that $g(w)$ is analytic in a neighborhood of the origin, $g(0) = 0$, and $g'(0) \neq 0$. Then g^{-1} can be expressed as the following power series near the origin:*

$$g^{-1}(z) = \sum_{n=1}^{\infty} \frac{z^n}{n!} \left[\frac{d^{n-1}}{dw^{n-1}} \left(\frac{w}{g(w)} \right)^n \right]_{w=0}.$$

Moreover, for any $\varphi : \mathbb{C} \rightarrow \mathbb{C}$ analytic in a neighborhood around the origin,

$$\varphi(g^{-1}(z)) = \varphi(0) + \sum_{n=1}^{\infty} \frac{z^n}{n!} \left[\frac{d^{n-1}}{dw^{n-1}} \left(\varphi'(w) \frac{w}{g(w)} \right)^n \right]_{w=0}.$$

Appendix B. Regularly varying functions

We define regularly varying functions and state Karamata’s theorems. See [4] for a rigorous treatment of the theory of regularly varying functions.

Definition B.1. A positive measurable function $f(x)$ is said to be regularly varying with index $\beta \in \mathbb{R}$ if

$$\lim_{x \rightarrow \infty} \frac{f(rx)}{f(x)} = r^\beta \quad \text{for all } r > 0.$$

Theorem B.1. (Karamata’s theorem, direct part [4].) *Let $f(x) : [a, \infty) \rightarrow [a, \infty)$ be a regularly varying function with index $\beta \in \mathbb{R}$. Then*

$$\lim_{x \rightarrow \infty} \frac{x^{\alpha+1}f(x)}{\int_a^x y^\alpha f(y) dy} = \alpha + \beta + 1 \quad \text{for all } \alpha > -\beta - 1$$

and

$$\lim_{x \rightarrow \infty} \frac{x^{\alpha+1}f(x)}{\int_x^\infty y^\alpha f(y) dy} = -(\alpha + \beta + 1) \quad \text{for all } \alpha < -\beta - 1.$$

We will use the following converse to the above Karamata’s theorem.

Theorem B.2. (Karamata’s theorem, converse [4].) *Let $f(x)$ be a positive, measurable, and locally integrable function on $[a, \infty)$ and $\beta \in \mathbb{R}$. Then the following hold:*

(a) *If there exists $\alpha > -\beta - 1$ such that*

$$\lim_{x \rightarrow \infty} \frac{x^{\alpha+1}f(x)}{\int_a^x y^\alpha f(y) dy} = \alpha + \beta + 1,$$

then $f(x)$ is a regularly varying function with index β .

(b) *If, for some $\alpha < -\beta - 1$,*

$$\lim_{x \rightarrow \infty} \frac{x^{\alpha+1}f(x)}{\int_x^\infty y^\alpha f(y) dy} = -(\alpha + \beta + 1),$$

then $f(x)$ is a regularly varying function with index β .

Appendix C. Proof of Lemma 2.5

Proof of Lemma 2.5. First we show that the tree $\mathcal{S}_r(\varphi, T)$ obtained by pruning a Galton–Watson tree $T \stackrel{d}{\sim} \mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$ is also distributed as a Galton–Watson tree.

For $T \stackrel{d}{\sim} \mu$ and $s \geq 0$, let $T|_{\leq s}$ denote a subtree of T consisting of all points x in the metric space T at distance no greater than s from the root ρ , i.e.,

$$T|_{\leq s} = \{x \in T : d(x, \rho) \leq s\}.$$

Let $T|_{=s}$ denote the points in T at distance s from the root ρ , i.e.,

$$T|_{=s} = \{x \in T : d(x, \rho) = s\}.$$

Let $\mathcal{F}_s^0 = \sigma(T|_{\leq s})$ be a sigma-algebra generated by the history up to time s (including branching history) of the Galton–Watson process that induces T . The future of the Galton–Watson process after time s consists of the descendant subtrees

$$\{\Delta_{x,T} : x \in T|_{=s}\}.$$

Let $\mathcal{F}'_s = \sigma(\Delta_{x,T} : x \in T|_{=s})$ be a sigma-algebra generated by the future events after time s . The measure μ being a Galton–Watson measure (i.e., $\mu \equiv \mathcal{GW}(\{q_k\}, \lambda)$) is equivalent to $T|_{=s}$ being a continuous-time Markov branching process (see [3, 13]). That is, there exists a filtration $\mathcal{F}_s \supset \mathcal{F}_s^0$ such that the following hold:

- 1. The Markov property is satisfied:

$$\mathbb{P}(A | \mathcal{F}_s) = \mathbb{P}(A | T|_{=s}) \quad \forall A \in \mathcal{F}'_s.$$

- 2. Conditioned on $T|_{=s}$, the subtrees $\{\Delta_{x,T} : x \in T|_{=s}\}$ of T , denoted here by

$$\left(\{\Delta_{x,T} : x \in T|_{=s}\} \middle| T|_{=s}\right),$$

are independent.

Next, let

$$\mathcal{I}_s = \sigma(\Delta_{x, \mathcal{S}_t(\varphi, T)} : x \in \mathcal{S}_t(\varphi, T)|_{=s})$$

be a sigma-algebra generated by the future events of $\mathcal{S}_t(\varphi, T)|_{=s}$, after time s . Then, since

$$\begin{aligned} \{\Delta_{x, \mathcal{S}_t(\varphi, T)} : x \in \mathcal{S}_t(\varphi, T)|_{=s}\} &= \{\mathcal{S}_t(\varphi, \Delta_{x,T}) : x \in \mathcal{S}_t(\varphi, T)|_{=s}\} \\ &= \{\mathcal{S}_t(\varphi, \Delta_{x,T}) : x \in T|_{=s} \text{ such that } \mathcal{S}_t(\varphi, \Delta_{x,T}) \neq \emptyset\}, \end{aligned}$$

we have

$$\mathcal{I}_s = \sigma(\mathcal{S}_t(\varphi, \Delta_{x,T}) : x \in T|_{=s}) \subset \mathcal{F}'_s.$$

We claim that conditioned on the event $\{\mathcal{S}_t(\varphi, T) \neq \emptyset\}$, the partition/annihilation evolution $\mathcal{S}_t(\varphi, T)|_{=s}$ is a continuous-time Markov branching process with respect to the filtration \mathcal{F}_s . Indeed, we have the following:

- 1. The Markov property is satisfied:

$$\mathbb{P}(A | \mathcal{F}_s) = \mathbb{P}(A | T|_{=s}) = \mathbb{P}(A | \mathcal{S}_t(\varphi, T)|_{=s}) \quad \forall A \in \mathcal{I}_s \subset \mathcal{F}'_s.$$

Let $\mathbb{P}_{\neq \emptyset}(A) = \mathbb{P}(A | \mathcal{S}_t(\varphi, T) \neq \emptyset)$. Then

$$\mathbb{P}_{\neq \emptyset}(A | \mathcal{F}_s) = \mathbb{P}_{\neq \emptyset}(A | \mathcal{S}_t(\varphi, T)|_{=s}) \quad \forall A \in \mathcal{I}_s.$$

2. Conditioned on $\mathcal{S}_t(\varphi, T)|_{=s}$, the subtrees

$$\{\Delta_{x, \mathcal{S}_t(\varphi, T)} : x \in \mathcal{S}_t(\varphi, T)|_{=s}\} = \{\mathcal{S}_t(\varphi, \Delta_{x, T}) : x \in T|_{=s}, \mathcal{S}_t(\varphi, \Delta_{x, T}) \neq \phi\}$$

of $\mathcal{S}_t(\varphi, T)$ are independent.

In order to characterize the dendritic structure of $\mathcal{S}_t(\varphi, T)$, we start an upward exploration from the root $\rho \in T$ and proceed to the nearest internal vertex v of T (i.e., $\text{par}(v) = \rho$). For a pair of integers $k \geq 2$ and $m \geq 0$, we have

$$\mathbb{P}(\text{br}_T(v) = k, \text{br}_{\mathcal{S}_t(\varphi, T)}(v) = m \mid \mathcal{S}_t(\varphi, T) \neq \phi) = \binom{k}{m} (1 - p_t)^{k-m} p_t^m \frac{q_k}{p_t}, \tag{49}$$

where $\text{br}_T(v)$ and $\text{br}_{\mathcal{S}_t(\varphi, T)}(v)$ denote the branching numbers of the vertex v in T and $\mathcal{S}_t(\varphi, T)$ respectively. Here, the event $\text{br}_{\mathcal{S}_t(\varphi, T)}(v) = 1$ corresponds to the case when the vertex v is removed via series reduction. Thus, the case $m = 1$ will be treated separately.

Next, we would like to find an expression for the branching probability g_m of a pruned tree $\mathcal{S}_t(\varphi, T)$. For a given integer $m \geq 2$,

$$\mathbb{P}(\text{br}_{\mathcal{S}_t(\varphi, T)}(v) = m \mid \mathcal{S}_t(\varphi, T) \neq \phi) = (1 - p_t)^{-m} p_t^{m-1} \sum_{k=m}^{\infty} \binom{k}{m} (1 - p_t)^k q_k.$$

Therefore, for $m \geq 2$,

$$\begin{aligned} g_m &= \mathbb{P}(\text{br}_{\mathcal{S}_t(\varphi, T)}(v) = m \mid \mathcal{S}_t(\varphi, T) \neq \phi, \text{br}_{\mathcal{S}_t(\varphi, T)}(v) \neq 1) \\ &= (1 - p_t)^{-m} p_t^{m-1} \frac{\sum_{k=m}^{\infty} \binom{k}{m} p^k q_k}{1 - (1 - p_t)^{-1} \sum_{k=2}^{\infty} k p^k q_k} = \frac{p_t^{m-1}}{m!} Q^{(m)}(1 - p_t) (1 - Q'(1 - p_t))^{-1}. \end{aligned}$$

The corresponding generating function of $\{g_k\}$ is equal to

$$\begin{aligned} G(z) &= \sum_{m=0}^{\infty} z^m g_m = g_0 + \frac{p_t^{-1}}{1 - (1 - p_t)^{-1} \sum_{k=2}^{\infty} k p^k q_k} \sum_{m=2}^{\infty} \sum_{k=m}^{\infty} (z(1 - p_t)^{-1} p_t)^m \binom{k}{m} p^k q_k \\ &= g_0 + \frac{p_t^{-1}}{1 - (1 - p_t)^{-1} \sum_{k=2}^{\infty} k p^k q_k} \sum_{k=2}^{\infty} \sum_{m=2}^k \binom{k}{m} (z(1 - p_t)^{-1} p_t)^m p^k q_k \\ &= g_0 + \frac{p_t^{-1}}{1 - Q'(1 - p_t)} (Q(z + (1 - z)(1 - p_t)) - Q(1 - p_t) - z p_t' Q(1 - p_t)) \\ &= z + g_0 + \frac{Q(z + (1 - z)(1 - p_t)) - Q(1 - p_t) - z p_t}{p_t(1 - Q'(1 - p_t))}, \tag{50} \end{aligned}$$

by the binomial theorem. Next, we plug $z = 1$ into (50), obtaining

$$g_0 = \frac{Q(1 - p_t) - (1 - p_t)}{p_t(1 - Q'(1 - p_t))},$$

as $G(1) = 1$. Hence, (50) can be rewritten as (21). We proceed by differentiating $\frac{d}{dz}$ in (21), obtaining

$$G'(z) = \frac{Q'(1 - p_t + zp_t) - Q'(1 - p_t)}{1 - Q'(1 - p_t)}. \quad (51)$$

An edge in $\mathcal{S}_t(\varphi, T)$ is a union of edges in the tree obtained after pruning T , but before the series reduction, separated by the degree-two vertices. The probability of a degree-two vertex after pruning (but before the series reduction) is

$$\sum_{k=2}^{\infty} q_k k p_t (1 - p_t)^{k-1} = p'_t Q(1 - p_t).$$

Hence, by Wald's equation, the edge lengths in $\mathcal{S}_t(\varphi, T)$ are independent exponential random variables, each with rate

$$\lambda(1 - Q'(1 - p_t)).$$

Finally, we observe that if $\mu(T)$ is critical, then $Q'(1) = 1$ and (51) imply $G'(1) = 1$. That is, $\nu(T | T \neq \emptyset)$ is a critical Galton–Watson measure. \square

Acknowledgement

We are thankful to Matthias Winkel for pointing out to us the connection between the generalized dynamical pruning and the hereditary reduction.

Funding information

This research is supported by NSF awards DMS-1412557 (Y. K.) and EAR-1723033 (I. Z.).

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

- [1] ABRAHAM, R. AND DELMAS, J. F. (2015). β -coalescents and stable Galton–Watson trees. *ALEA Lat. Am. J. Prob. Math. Statist.* **12**, 451–476.
- [2] ABRAMOWITZ, M. AND STEGUN, I. A. (eds) (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, Washington, DC.
- [3] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, Berlin, Heidelberg.
- [4] BINGHAM, N. H., GOLDIE, C. M. AND TEUGELS, J. L. (1989). *Regular Variation*. Cambridge University Press.
- [5] BURD, G. A., WAYMIRE, E. C. AND WINN, R. D. (2000). A self-similar invariance of critical binary Galton–Watson trees. *Bernoulli* **6**, 1–21.
- [6] DEVROYE, L. AND KRUSZEWSKI, P. (1994). A note on the Horton–Strahler number for random trees. *Inf. Process. Lett.* **56**, 95–99.
- [7] DUQUESNE, T. AND LE GALL, J.-F. (2002). *Random Trees, Lévy Processes and Spatial Branching Processes* (Astérisque 281). Société Mathématique de France, Paris.
- [8] DUQUESNE, T. AND WINKEL, M. (2007). Growth of Lévy trees. *Prob. Theory Relat. Fields* **139**, 313–371.
- [9] DUQUESNE, T. AND WINKEL, M. (2019). Hereditary tree growth and Lévy forests. *Stoch. Process. Appl.* **129**, 3690–3747.
- [10] EVANS, S. N. (2008). *Probability and Real Trees: École d'Été de Probabilités de Saint-Flour XXXV-2005*. Springer, Berlin, Heidelberg.
- [11] EVANS, S. N., PITMAN, J. AND WINTER, A. (2006). Rayleigh processes, real trees, and root growth with re-grafting. *Prob. Theory Relat. Fields* **134**, 81–126.

- [12] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd edn. John Wiley, New York.
- [13] HARRIS, T. E. (1963). *The Theory of Branching Processes*. Springer, Berlin, Heidelberg.
- [14] HE, H. AND WINKEL, M. (2014). *Invariance principles for pruning processes of Galton–Watson trees*. Preprint. Available at <https://arxiv.org/abs/1409.1014>.
- [15] HORTON, R. E. (1945). Erosional development of streams and their drainage basins: hydrophysical approach to quantitative morphology. *Geol. Soc. Amer. Bull.* **56**, 275–370.
- [16] KESTEN, H. (1987). Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Prob. Statist.* **22**, 425–487.
- [17] KOVCHEGOV, Y. AND ZALIAPIN, I. (2016). Horton law in self-similar trees. *Fractals* **24**, article no. 1650017.
- [18] KOVCHEGOV, Y. AND ZALIAPIN, I. (2019). Random self-similar trees and a hierarchical branching process. *Stoch. Process. Appl.* **129**, 2528–2560.
- [19] KOVCHEGOV, Y. AND ZALIAPIN, I. (2020). Dynamical pruning of rooted trees with applications to 1D ballistic annihilation. *J. Statist. Phys.* **181**, 618–672.
- [20] KOVCHEGOV, Y. AND ZALIAPIN, I. (2020). Random self-similar trees: A mathematical theory of Horton laws. *Prob. Surveys* **17**, 1–213.
- [21] KOVCHEGOV, Y. AND ZALIAPIN, I. (2021). Invariance and attraction properties of Galton–Watson trees. *Bernoulli* **27**, 1789–1823.
- [22] KOVCHEGOV, Y., ZALIAPIN, I. AND BEN-ZION, Y. (2022) Invariant Galton–Watson branching process for earthquake occurrence. *Geophys. J. Internat.* **231**, 567–583.
- [23] LE JAN, Y. (1991). Superprocesses and projective limits of branching Markov process. *Ann. Inst. H. Poincaré Prob. Statist.* **27**, 91–106.
- [24] MCCONNELL, M. AND GUPTA, V. (2008). A proof of the Horton law of stream numbers for the Tokunaga model of river networks. *Fractals* **16**, 227–233.
- [25] NEVEU, J. (1986). Erasing a branching tree. *Adv. Appl. Prob.* **1**, 101–108.
- [26] NEVEU, J. AND PITMAN, J. (1989). Renewal property of the extrema and tree property of the excursion of a one-dimensional Brownian motion. In *Séminaire de Probabilités XXIII*, Springer, Berlin, Heidelberg, pp. 239–247.
- [27] NEVEU, J. AND PITMAN, J. (1989). The branching process in a Brownian excursion. In *Séminaire de Probabilités XXIII*, Springer, Berlin, Heidelberg, pp. 248–257.
- [28] NEWMAN, W. I., TURCOTTE, D. L. AND GABRIELOV, A. M. (1997). Fractal trees with side branching. *Fractals* **5**, 603–614.
- [29] PECKHAM, S. D. (1995). New results for self-similar trees with applications to river networks. *Water Resources Res.* **31**, 1023–1029.
- [30] STRAHLER, A. N. (1957). Quantitative analysis of watershed geomorphology. *Trans. Am. Geophys. Union* **38**, 913–920.
- [31] WHITTAKER, E. T. AND WATSON, G. N. (1927). *A Course of Modern Analysis*, 4th edn. Cambridge University Press.
- [32] ZOLOTAREV, V. M. (1957). More exact statements of several theorems in the theory of branching processes. *Theory Prob. Appl.* **2**, 256–266.